A Study on Approximation of Funcions

Dissertation submitted to

Visvesvaraya National Institute of Technology, Nagpur

in partial fulfillment of the requirement for the award of the degree

Master of Science

in

Mathematics

by

Riya Khedekar

(MS23MTH002)

under the supervision of

Dr. Vijender Nallapu

Department of Mathematics

Visvesvaraya National Institute of Technology

Nagpur - 440010 (India)

June 2025

A Study on Approximation of Funcions

Dissertation submitted to

Visvesvaraya National Institute of Technology, Nagpur

in partial fulfillment of the requirement for the award of the degree

Master of Science

in

Mathematics

by

Riya Khedekar

(MS23MTH002)

under the supervision of

Dr. Vijender Nallapu

Department of Mathematics

Visvesvaraya National Institute of Technology

Nagpur - 440010 (India)

June 2025

©Visvesvaraya National Institute of Technology(VNIT) 2025

Department of Mathematics

Visvesvaraya National Institute of Technology, Nagpur

Declaration

I, Riya Khedekar, hereby declare that this dissertation titled "**A Study on Approximation of Functions**" is carried out by me in the Department of Mathematics of Visvesvaraya National Institute of Technology, Nagpur. The work is original and has not been submitted earlier whole or in part for the award of any degree/diploma at this or any other Institution/University.

Riya Khedekar

Date:

Certificate

This is to certify that the dissertation titled " **A Study on Approximation of Functions**" is submitted by **Riya Khedekar** in partial fulfillment of the requirements for the award of the degree of **Master of Science in Mathematics**, VNIT Nagpur. The work is comprehensive,complete and fit for final evaluation.

Dr. Vijender Nallapu

Assistant Professor, Mathematics

VNIT, Nagpur

Head, Department of Mathematics

VNIT, Nagpur

Date:

VISVESVARAYA NATIONAL INSTITUTE OF TECHNOLOGY

NAGPUR

Non-Plagiarism Certificate

Certified that the M.Sc. dissertation titled **"A Study on Approximation of Functions"** submitted by **Riya Khedekar**, Enrollment No. **MS23MTH002**, ID No. **30078**, is original in nature and not copied from any literature material. Due credit has been given to the original researcher wherever referred.

A summary of the plagiarism report using turnitin software is given below:

Signature :

Student Name : Riya Khedekar

Enrollment No. : MS23MTH002

ID No. 30078

Dr. Vijender Nallapu

Assistant Professor, Mathematics

VNIT, Nagpur

Acknowledgment

It is with immense pleasure and heartfelt gratitude that I acknowledge the invaluable support of my supervisor, **Dr. Vijender Nallapu**, Assistant Professor, Department of Mathematics, Visvesvaraya National Institute of Technology (VNIT), Nagpur. His expert guidance, insightful suggestions, and constant encouragement were instrumental in the successful completion of this project.

I extend my sincere thanks to **Dr. Pallavi Mahale**, Professor and Head of the Department of Mathematics at VNIT Nagpur, for providing an excellent academic environment and for her continuous support throughout the project period. I am also grateful to the project progress and evaluation committee members, **Dr. Naga Raju Gande** , Associate Professor and **Dr. Sourav Bhattacharya**, Assistant Professor in the Department, for their valuable feedback and encouragement during various stages of the project.

My sincere appreciation goes to all the faculty members of the Department of Mathematics, VNIT Nagpur, for their constant support and guidance.

I would like to express my special thanks to Ms. Richa Jain and Mr. Soumya Ghosh, Ph.D. Research Scholars, for their generous help, insightful discussions, and consistent encouragement. Their commitment and passion have inspired me throughout this journey.

Lastly, I am deeply grateful to my parents and family members for their unconditional love, unwavering support, and constant motivation, which have been a pillar of strength throughout this endeavor.

Riya Khedekar

Abstract

This thesis is devoted to study a few foremost existing results of approximation theory. In Chapter 1, we record a few definition and important results required for our study. In Chapter 2, the celebrated Weierstrass approximation theorem with existence type proof due to Weierstrass and constructive type proof due to Bernstein is discussed. Chapter 3 is devoted to the study of Bohman–Korovkin theorem and its application. In Chapter 4, we discuss the existence and uniqueness of best approximation to a continuous function defined on a compact interval on the real line.

List of Notations

$\mathbb{R}$	: Set of real numbers
$\mathbb{N}$	: Set of natural numbers
inf	: Infimum (or greatest lower bound)
sup	: Supremum (or least upper bound)
max	: Maximum
min	: Minimum
ϵ	: Belongs to
$\exists$	: There exist
$\forall$	: For all
$C[a, b]$	: Space of all continous functions on $[a, b]$

Contents

List of Notations **ii**

1 Introduction **1**

1.1 Basic definitions 1

2 Weierstrass Approximation Theorem **4**

2.1 Introduction 4

2.2 Weierstrass approximation theorem 4

2.3 Constructive proof of Weierstrass approximation theorem given by Bernstein . . 6

2.3.1 Bernstein proof of theorem 2.2.1 8

2.4 Trigonometric polynomials and Weierstrass second theorem 9

3 The Bohman–Korovkin Theorem **13**

3.1 Introduction 13

3.2 Statement of the Bohman–Korovkin theorem 13

3.3 An application of the Bohman–Korovkin theorem 16

4 Characterization of Best Approximation and Chebyshev Approximation **17**

4.1 Introduction 17

4.2 Best approximation in the uniform norm 17

4.3 Chebyshev's problem and error symmetry lemma 18

4.4 Existence of an alternating set 19

4.5 Sufficiency of alternating set 20

4.6 Uniqueness of best approximation 21

4.7 Example: Best approximation of sin(4x) 21

4.8 Chebyshev's approximation problem 22

Conclusion **30**

References **31**

Chapter 1

Introduction

One of the central themes in approximation theory is to represent an arbitrary function in terms of other functions which are nicer or simpler or both. In many science and engineering application, functions that model natural phenomena are highly intricate and difficult to analyze directly. Replacing such functions with polynomials or other elementary functions allows us to perform computations more efficiently. The theory of approximation of functions of a real variable had its origin in the work of Chebyshev and the well-known thoerem of Weierstrass on the approximations of continous functions by polynomials.
Before exploring the main results, let us recall few important definitions and results.

1.1 Basic definitions

Definition 1.1.1. *Let P denote a vector space over the field $\mathbb{R}$ or $\mathbb{C}$. The function $\|\cdot\| : P \to \mathbb{R}$ serves the role of a norm if it assigns a non-negative real value to each element of P and the following conditions hold true for all vectors $x, y \in P$ and any scalar α:*

1. ***Positivity and zero condition:***
   The norm of a vector is non-negative, and $\|x\| = 0$ holds exclusively when $x = 0$.
2. ***Homogeneity:***
   For any scalar α, the norm satisfies $\|\alpha x\| = |\alpha| \cdot \|x\|$.
3. ***Triangle Inequality:***
   The computed norm of a vector that arises from adding two vectors together does not exceed the sum of their norms:
   $\|x + y\| \leq \|x\| + \|y\|$.

Equipping a vector space P with a norm transforms it into what is formally described as normed linear space.

Example 1.1.1. *Consider the space $\mathbb{R}^n$ with the Euclidean norm $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$, which is a normed linear space.*

Definition 1.1.2. *The set **C[a, b]** represents the collection of all continuous functions mapping from the interval $[a, b]$ to the real numbers, equipped with the supremum norm.*

Definition 1.1.3. *Suppose y is a function whose domain is a closed interval $[p, q]$, the **supremum (or uniform) norm** is defined by:*

$$\|f\|_\infty = \sup_{r \in [p,q]} |y(r)|.$$

Definition 1.1.4. *Let $D \subseteq \mathbb{R}$. A **sequence of functions** refers to a collection $\{f_n\}_{n=1}^{\infty}$ where for every n, the function f_n takes inputs from D and produces outputs in $\mathbb{R}$ or $\mathbb{C}$.For any x that is held constant in D, the values $\{f_n(x)\}$ form a numerical sequence of real or complex values. Our focus is on how this sequence behaves as $n \to \infty$, either pointwise or uniformly.*

Definition 1.1.5. *A function sequence $\{d_n\}$ from the space $C[b, c]$ is said to converge uniformly to a function $d \in C[b, c]$ if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,*

$$\sup_{x \in [b,c]} |d_n(x) - d(x)| < \varepsilon.$$

Definition 1.1.6. *A function $g : [p, q] \to \mathbb{R}$ is said to be uniformly continuous on the interval $[p, q]$ if, for every $\eta > 0$, there exists a $\theta > 0$ such that for all $u, v \in [p, q]$, the condition*

$$|u - v| < \theta$$

implies that

$$|g(u) - g(v)| < \eta.$$

Definition 1.1.7. *Consider a bounded function $l\colon [p, q] \to \mathbb{R}$. The modulus of continuity of l, denoted by ω_l, is given by: For any $\delta > 0$,*

$$\omega_l(\delta) = \sup \{|l(x) - l(y)| : x, y \in [p, q],\ |x - y| \leq \delta\}$$

Definition 1.1.8. *The space $\mathbb{P}_n$ is a collection of all real polynomials with a degree up to n. That is:*

$$\mathbb{P}_n = \left\{ p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mid a_0, a_1, \ldots, a_n \in \mathbb{R} \right\}.$$

Definition 1.1.9. *Let $g \in C[0,1]$ be a continuous real-valued function on $[0,1]$. The **Bernstein polynomial** whose degree does not exceed n is expressed as :*

$$B_n(g;r) = \sum_{k=0}^{n} g\left(\frac{k}{n}\right) \binom{n}{k} r^k (1-r)^{n-k}, \quad \textit{for } r \in [0,1].$$

Definition 1.1.10. *Let $T : C[a,b] \to C[a,b]$ be a mapping . The mapping T is termed a **positive linear operator** if it meets these two conditions:*

- ***Linearity:** For any two functions $f, g \in C[a,b]$ and scalars $\alpha, \beta \in \mathbb{R}$,*

$$T(\alpha f + \beta g) = \alpha T(f) + \beta T(g).$$

- ***Positivity:** If $f(x) \geq 0$ for every $x \in [a,b]$, then*

$$T(f)(x) \geq 0 \quad \textit{for each } x \in [a,b].$$

Definition 1.1.11. *In the context of Korovkin-type approximation theorems, a specific set of basic functions is commonly used to assess uniform convergence. These are recognized as the **Korovkin test functions**, and they are defined as follows:*

$$f_0(x) = 1, \quad f_1(x) = x, \quad f_2(x) = x^2.$$

Definition 1.1.12. *The first-kind Chebyshev polynomial whose degree is n, denoted by $T_n(x)$, is given on $[-1,1]$ as:*

$$T_n(x) = \cos(n \cdot \arccos x).$$

Definition 1.1.13. *Let $\mathbf{C}_{2\pi}$ denote the space of all continuous functions that are periodic with period 2π. The norm in this space is given by*

$$\|f\| = \sup_{x \in \mathbb{R}} |f(x)|.$$

Chapter 2

Weierstrass Approximation Theorem

2.1 Introduction

In 1885, Karl Weierstrass proved a remarkable result that extended the scope of approximation. This result, now known as the Weierstrass Approximation Theorem, established a foundation for modern approximation theory.

In this chapter, we discuss Weierstrass Approximation Theorem.

2.2 Weierstrass approximation theorem

Theorem 2.2.1. *[1] Let f be a continous function on $[0,1]$. For any positive number ε, there exists a polynomial p such that*

$$\|f - p\|_\infty < \varepsilon,$$

where

$$\|f - p\|_\infty = \sup_{x\in[0,1]} |f(x) - p(x)|$$

Proof. Assume $f \in C[0,1]$. Since f is continuous on the compact interval $[0,1]$, it is uniformly continuous and bounded. Define $M = \sup_{x\in[0,1]} |f(x)|$. We extend f to the real line by setting $f(x) = 0$ for all $x \notin [0,1]$. The extended function remains uniformly continuous on $\mathbb{R}$ and bounded. Now define a family of polynomials: $B_n(x) = C_n(1-x^2)^n$, for $|x| \leq 1$, where the constant C_n is chosen such that

$$\int_{-1}^{1} B_n(x)\,dx = 1.$$

It can be shown that $C_n \sim \sqrt{n}$, and for any fixed $\delta > 0$,

$$B_n(x) \leq \sqrt{n}(1-\delta^2)^n \quad \text{whenever } \delta \leq |x| \leq 1.$$

Hence, $B_n(x) \to 0$ uniformly on the complement of $(-\delta, \delta) \subset [-1, 1]$. Now define a sequence of approximating functions for $x \in [0, 1]$:

$$P_n(x) = \int_{-1}^{1} f(x+t)B_n(t)\,dt$$

Since $f(x+t) = 0$ outside $[0, 1]$, the integral becomes:

$$P_n(x) = \int_{0}^{1} f(y)B_n(y-x)\,dy,$$

after the substitution $y = x + t$.

We will now prove that $P_n(x) \to f(x)$ uniformly on $[0, 1]$. Let $\varepsilon > 0$. By the uniform continuity of f, there exists $\delta > 0$ such that

$$|f(t_1) - f(t_2)| < \varepsilon \quad \text{whenever } |t_1 - t_2| < \delta.$$

Then,

$$|P_n(x) - f(x)| = \left| \int_{-1}^{1} (f(x+t) - f(x))\, B_n(t)\,dt \right|$$

It can be written that

$$\left| \int_{-1}^{1} (f(x+t) - f(x))\, B_n(t)\,dt \right| \leq \left| \int_{|t|<\delta} (f(x+t) - f(x))B_n(t)\,dt \right| + \left| \int_{|t|\geq\delta} (f(x+t) - f(x))B_n(t)\,dt \right|$$

Since $|f(x+t) - f(x)| < \varepsilon$ for $|t| < \delta$, it follows that

$$\int_{|t|<\delta} |f(x+t) - f(x)|B_n(t)\,dt < \varepsilon \int_{-1}^{1} B_n(t)\,dt = \varepsilon.$$

Since $|f(t)| \leq M$ for all $|t| \geq \delta$, we get

$$|f(x+t) - f(x)| \leq 2M \Rightarrow \int_{|t|\geq\delta} |f(x+t) - f(x)|B_n(t)\,dt \leq 2M \int_{|t|\geq\delta} B_n(t)\,dt.$$

Since $B_n(t) \to 0$ as $n \to \infty$ uniformly outside $(-\delta, \delta)$ we get $\int_{|t|\geq\delta} B_n(t)\, dt \to 0$ as $n \to \infty$. Hence, for sufficiently large n, this term is also less than ε. Combining both estimates:

$$|P_n(x) - f(x)| < \varepsilon + \varepsilon = 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we conclude that:

$$\lim_{n\to\infty} \|P_n - f\|_\infty = 0.$$

Thus, the sequence $\{P_n(x)\}$ consists of polynomials that converge uniformly to f on $[0, 1]$. This proves the Weierstrass approximation theorem. □

2.3 Constructive proof of Weierstrass approximation theorem given by Bernstein

To demonstrate the Weierstrass approximation theorem in a constructive manner, we utilize Bernstein polynomial which are tailored to achieve uniform convergence to any continous function defined on $[0, 1]$.

Consider a continuous function h on the interval $[0, 1]$. Its Bernstein polynomial of degree n is given by

$$B_n(h; x) = \sum_{k=0}^{n} h\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}, \quad \text{for } x \in [0, 1].$$

Let us verify the behavior of Bernstein polynomials on simple functions to build the foundation for the proof.

Lemma 2.3.1. *As illustrative cases, we take three basic functions: the constant function $h_0(t) = 1$, the identity function $h_1(t) = t$, and the quadratic function $h_2(t) = t^2$. Then:*

(i) $B_n(h_0)(t) = 1$, *and* $B_n(h_1)(t) = t$.

(ii) $B_n(h_2)(t) = t^2 + \frac{1}{n}t(1-t)$. *Thus,* $B_n(h_2) \to h_2$ *uniformly as* $n \to \infty$.

(iii) For $0 \leq t \leq 1$,

$$\sum_{k=0}^{n} \left(\frac{k}{n} - t\right)^2 \binom{n}{k} t^k (1-t)^{n-k} = \frac{t(1-t)}{n} \leq \frac{1}{4n}.$$

(iv) Let $\delta > 0$ and define $F := \{k \in \{0, 1, \ldots, n\} : \left|\frac{k}{n} - t\right| \geq \delta\}$. Then:

$$\sum_{k \in F} \binom{n}{k} t^k (1-t)^{n-k} \leq \frac{1}{4n\delta^2}.$$

Proof. **(i)** For $h_0(t) = 1$, we have:

$$B_n(h_0)(t) = \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} = (t + (1-t))^n = 1.$$

For $h_1(t) = t$, observe:

$$B_n(h_1)(t) = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} t^k (1-t)^{n-k} = t.$$

(ii) Similarly, we can show:

$$B_n(h_2)(t) = t^2 + \frac{1}{n} t(1-t).$$

Hence, the error term:

$$\|B_n(h_2) - h_2\| = \left\| \frac{1}{n} t(1-t) \right\| \leq \frac{1}{4n} \to 0 \quad \text{as } n \to \infty.$$

(iii) By expanding the square:

$$\left(\frac{k}{n} - t\right)^2 = \left(\frac{k}{n}\right)^2 - 2t \cdot \frac{k}{n} + t^2.$$

Using linearity and known expressions for $B_n(h_1)$ and $B_n(h_2)$, the result follows:

$$\sum_{k=0}^{n} \left(\frac{k}{n} - t\right)^2 \binom{n}{k} t^k (1-t)^{n-k} = \frac{t(1-t)}{n}.$$

(iv) Since $\left|\frac{k}{n} - t\right| \geq \delta$ for $k \in F$, we get:

$$\sum_{k \in F} \binom{n}{k} t^k (1-t)^{n-k} \leq \frac{1}{\delta^2} \sum_{k=0}^{n} \left(\frac{k}{n} - t\right)^2 \binom{n}{k} t^k (1-t)^{n-k} \leq \frac{1}{4n\delta^2}.$$

□

2.3.1 Bernstein proof of theorem 2.2.1

Proof. [2] Let h be a continuous function on $[0,1]$, and fix $\varepsilon > 0$ be arbitrary. By uniform continuity of h on $[0,1]$, for any $\varepsilon > 0$, a $\delta > 0$ exists such that

$$|t - y| < \delta \Rightarrow |h(t) - h(y)| < \frac{\varepsilon}{2}.$$

We now estimate the error $\|h - B_n(h)\|$, where $B_n(h)$ denotes the n-th Bernstein polynomial of h. Recall that

$$B_n(h)(t) = \sum_{k=0}^{n} h\left(\frac{k}{n}\right)\binom{n}{k} t^k (1-t)^{n-k}.$$

Then,

$$\begin{aligned} |h(t) - B_n(h)(t)| &= \left| \sum_{k=0}^{n} \left(h(t) - h\left(\frac{k}{n}\right)\right)\binom{n}{k} t^k (1-t)^{n-k} \right| \\ &\leq \sum_{k=0}^{n} \left| h(t) - h\left(\frac{k}{n}\right)\right| \binom{n}{k} t^k (1-t)^{n-k}. \end{aligned}$$

Now let n be a fixed natural number, and define the collection

$$F = \left\{ k \in \{0, 1, \ldots, n\} \mid \left| \frac{k}{n} - t \right| \geq \delta \right\}.$$

The sum can be broken down accordingly:

- For $k \notin F$, i.e., $\left|\frac{k}{n} - t\right| < \delta$, we have by uniform continuity:

$$\left| h(t) - h\left(\frac{k}{n}\right)\right| < \frac{\varepsilon}{2}.$$

- For $k \in F$, since h is bounded on $[0,1]$, say $\|h\| = \sup_{t\in[0,1]} |h(t)|$, we have:

$$\left| h(t) - h\left(\frac{k}{n}\right)\right| \leq 2\|h\|.$$

Consequently, we write:

$$|h(t) - B_n(h)(t)| \leq \frac{\varepsilon}{2} \sum_{k\notin F} \binom{n}{k} t^k (1-t)^{n-k} + 2\|f\| \sum_{k\in F} \binom{n}{k} t^k (1-t)^{n-k}.$$

Applying the lemma (iv), we obtain:

$$\sum_{k\in F}\binom{n}{k}t^k(1-t)^{n-k} < \frac{1}{4n\delta^2}.$$

Therefore,

$$|h(t) - B_n(h)(t)| < \frac{\varepsilon}{2} + 2\|h\| \cdot \frac{1}{4n\delta^2}.$$

Achieving a bound less than ε on the right requires choosing n such that

$$2\|h\| \cdot \frac{1}{4n\delta^2} < \frac{\varepsilon}{2}, \quad \text{or equivalently,} \quad n > \frac{\|h\|}{\varepsilon\delta^2}.$$

Thus, for sufficiently large n, we have $\|h - B_n(h)\| < \varepsilon$, completing the proof. □

2.4 Trigonometric polynomials and Weierstrass second theorem

The classical Weierstrass Theorem ensures uniform approximation of continuous functions on $[a, b]$ by polynomials. The Weierstrass second theorem ensures that periodic functions can be approximated by trignometric functions.

Trigonometric polynomials

We begin with the definition of trigonometric polynomials and the space in which they are studied.

Definition 2.4.1. *A trigonometric polynomial of degree at most n is given by*

$$T_n(x) = a_0 + \sum_{k=1}^{n} \left(a_k \cos(kx) + b_k \sin(kx)\right),$$

where $a_0, a_1, \ldots, a_n$, $b_1, \ldots, b_n$ are real numbers. Let $\mathcal{T}_n$ be the set of trigonometric polynomials of degree at most n. Then, the complete set of trigonometric polynomials is given by

$$\mathcal{T} = \bigcup_{n=0}^{\infty} \mathcal{T}_n.$$

Let $C_{2\pi}$ denote the set of all continuous functions on $\mathbb{R}$ that are 2π-periodic, endowed with the

uniform norm:

$$\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)| = \max_{x\in[0,2\pi]} |f(x)|.$$

Trigonometric polynomials serve as natural approximants for periodic functions due to their periodicity and smooth structure.

Best approximation by trigonometric polynomials

Every function $f \in C_{2\pi}$ has a best approximation from $\mathcal{T}_n$, and when f is even, this approximant can also be chosen to be even.

Corollary 2.4.1. *For any function $f \in C_{2\pi}$ and $n \in \mathbb{N}$, there exists a trigonometric polynomial $T \in \mathcal{T}_n$ satisfying*

$$\|f - T\|_\infty = \inf_{S\in\mathcal{T}_n} \|f - S\|_\infty.$$

Furthermore, if f is even, then T can be chosen to be even as well.

Proof. Since $\mathcal{T}_n$ constitutes a finite-dimensional subspace of the space $C_{2\pi}$ and the norm is continuous, the best approximation exists by standard results from analysis.

Now suppose f is even. Let $T^* \in \mathcal{T}_n$ be a best approximation. Define $S(x) = T^*(-x)$. As trigonometric polynomials are closed under reflection, $S \in \mathcal{T}_n$, and since f is even, we get

$$|f(x) - S(x)| = |f(-x) - T^*(x)| = |f(x) - T^*(x)|,$$

so $\|f - S\|_\infty = \|f - T^*\|_\infty$.

Now define $T(x) = \frac{1}{2}\left(T^*(x) + S(x)\right) = \frac{1}{2}\left(T^*(x) + T^*(-x)\right)$ which is even. Then

$$\|f - T\|_\infty \le \frac{1}{2}\|f - T^*\|_\infty + \frac{1}{2}\|f - S\|_\infty = \|f - T^*\|_\infty.$$

Hence, T is an even trigonometric polynomial of best approximation. □

Now, we discuss Weierstrass second approximation theorem.

Theorem 2.4.1 (Weierstrass second approximation theorem)**.** *[2] Let $f \in C_{2\pi}$ and $\varepsilon > 0$. Then there exists a trigonometric polynomial $T \in \mathcal{T}$ such that*

$$\|f - T\|_\infty < \varepsilon.$$

Proof. We first prove for even functions and then using it to prove for general function.
Suppose $f \in C_{2\pi}$ is even. Consider the substitution $g(y) = f(\arccos y)$ on $y \in [-1, 1]$. Since $\arccos y$ is continuous and f is continuous and even, $g \in C[-1, 1]$. By the classical Weierstrass theorem, there exists a polynomial p such that

$$\sup_{y\in[-1,1]} |g(y) - p(y)| < \varepsilon.$$

Define $T(x) = p(\cos x)$. This is an even trigonometric polynomial, and for all $x \in [0, 2\pi]$,

$$|f(x) - T(x)| = |f(\arccos(\cos x)) - p(\cos x)| < \varepsilon,$$

so $\|f - T\|_\infty < \varepsilon$.
Let $f \in C_{2\pi}$. Define the auxiliary functions:

$$f_1(x) = f(x) + f(-x), \quad f_2(x) = [f(x) - f(-x)] \cdot \sin x.$$

Note that both f_1 and f_2 are even functions:

$$f_1(-x) = f(-x) + f(x) = f_1(x), \quad f_2(-x) = [f(-x) - f(x)] \cdot \sin(-x) = f_2(x).$$

Since f_1 and f_2 are even and continuous, there exist even trigonometric polynomials T_1 and T_2 such that

$$\|f_1 - T_1\|_\infty < \varepsilon \quad \text{and} \quad \|f_2 - T_2\|_\infty < \varepsilon.$$

Now consider the expression $2f(x)\sin^2 x = f_1(x)\sin^2 x + f_2(x)\sin x$. We approximate this as

$$2f(x)\sin^2 x \approx T_1(x)\sin^2 x + T_2(x)\sin x + 2\varepsilon.$$

Define the trigonometric polynomial:

$$T_3(x) := T_1(x)\sin^2 x + T_2(x)\sin x.$$

Then we obtain:

$$f(x) \approx T_3(x) + 2\varepsilon.$$

To account for cosine components, consider the translated function $f\left(x - \frac{\pi}{2}\right)$, and apply the same

decomposition as in Step 2. Using a similar construction, one obtains an approximation:

$$2f(x)\cos^2 x \approx T_5(x) + \varepsilon,$$

where $T_5(x)$ is a trigonometric polynomial derived in the same way as $T_3(x)$. Now use the identity $\sin^2 x + \cos^2 x = 1$ to combine the two expressions:

$$f(x) = f(x)(\sin^2 x + \cos^2 x) = \frac{1}{2}\left(2f(x)\sin^2 x + 2f(x)\cos^2 x\right).$$

Substituting the approximations, define:

$$T(x) = \frac{1}{2}\left(T_3(x) + T_5(x)\right).$$

We get,

$$\|f - T\|_\infty < 2\varepsilon.$$

As ε was chosen arbitrarily, we conclude that f can be uniformly approximated by trigonometric polynomial. □

Chapter 3

The Bohman–Korovkin Theorem

3.1 Introduction

In the previous chapters, we studied how continuous functions on compact intervals can be approximated uniformly using polynomials and trigonometric polynomials. A key observation from Bernstein's approach was that approximating a small set of basic functions—such as 1, x, and x^2—appeared sufficient to guarantee uniform approximation for all continuous functions.

This phenomenon is not coincidental. The Bohman–Korovkin theorem provides a clear and powerful criterion for uniform convergence of sequences of positive linear mappings. Simply put, if the sequence converges on three specific test functions, it converges uniformly on all functions in $C[a, b]$.

The Bohman-Korovkin theorem greatly simplifies convergence proofs and has widespread applications in numerical analysis, functional approximation, and the theory of operators. In this chapter, we discuss Bohman–Korovkin theorem.

3.2 Statement of the Bohman–Korovkin theorem

Theorem 3.2.1 (Bohman–Korovkin theorem)**.** *[3] Suppose $D \subset \mathbb{R}$ is a compact set, and let $\{\Psi_m\}$ be a sequence of positive linear mappings operating on $C(D)$, the space of real-valued continuous functions defined on D, endowed with the supremum norm $\|\cdot\|_\infty$.*

Assume that the following condition holds for the test functions $g_0(x) = 1$, $g_1(x) = x$, and $g_2(x) = x^2$:

$$\lim_{m\to\infty} \|\Psi_m(g_i) - g_i\|_\infty = 0, \quad \textit{for each } i = 0, 1, 2.$$

Then, for any $g \in C(D)$, we have:

$$\lim_{m\to\infty} \|\Psi_m(g) - g\|_\infty = 0.$$

Proof. Consider a function $g \in C(D)$ and a fixed $\epsilon > 0$. Our objective is to show there exists a positive integer M such that for all $m \geq M$, the following holds:

$$\|\Psi_m(g) - g\|_\infty < \epsilon.$$

We start by defining the error term:

$$\Gamma_m(g)(t) := |\Psi_m(g)(t) - g(t)|,$$

and aim to bound this uniformly over all $t \in D$. Since Ψ_m is a positive linear mapping and $g \in C(D)$, we can write:

$$\Gamma_m(g)(t) = |\Psi_m(g)(t) - \Psi_m(1)(t) \cdot g(t) + \Psi_m(1)(t) \cdot g(t) - g(t)|.$$

Using the triangle inequality, this expression can be written as

$$\Gamma_m(g)(t) \leq |\Psi_m(g)(t) - \Psi_m(1)(t) \cdot g(t)| + |\Psi_m(1)(t) \cdot g(t) - g(t)|.$$

This simplifies to:

$$\Gamma_m(g)(t) \leq \Psi_m(|g(s) - g(t)|)(t) + \|g\|_\infty \cdot |\Psi_m(1)(t) - 1|.$$

Since $\|\Psi_m(1) - 1\|_\infty \to 0$ as $m \to \infty$, there exists $m_1 \in \mathbb{N}$ such that for all $m \geq m_1$,

$$\|g\|_\infty \cdot \|\Psi_m(1) - 1\|_\infty < \frac{\epsilon}{4}.$$

So,

$$\Gamma_m(g)(t) \leq \Psi_m(|g(s) - g(t)|)(t) + \frac{\epsilon}{4}.$$

Since g is continuous on the compact set D, it is uniformly continuous. Hence, for any $\epsilon > 0$,

there exists $\delta_1 > 0$ such that:

$$|t - s| < \delta_1 \quad \Longrightarrow \quad |g(t) - g(s)| < \epsilon_1$$

Subsequently,

$$|g(t) - g(s)| \leq \epsilon_1 + \frac{2\|g\|_\infty}{\delta_1^2}(t - s)^2,$$

which holds uniformly for all $t, s \in D$. Using the positivity and linearity of Ψ_m, we get

$$\Psi_m(|g(s) - g(t)|)(t) \leq \epsilon_1 \cdot \Psi_m(1)(t) + \frac{2\|g\|_\infty}{\delta_1^2} \cdot \Psi_m\left((s - t)^2\right)(t).$$

Thus

$$\Psi_m((s - t)^2)(t) = \Psi_m(s^2)(t) - 2t \cdot \Psi_m(s)(t) + t^2 \cdot \Psi_m(1)(t).$$

By the Korovkin condition, we have:

$$\|\Psi_m(s) - s\|_\infty \to 0 \quad \text{and} \quad \|\Psi_m(s^2) - s^2\|_\infty \to 0.$$

Hence, there exists $m_2 \in \mathbb{N}$ such that for all $m \geq m_2$,

$$|\Psi_m(s^2)(t) - t^2| < \theta, \quad |\Psi_m(s)(t) - t| < \theta, \quad |\Psi_m(1)(t) - 1| < \theta,$$

where $\theta > 0$ is small enough to meet the error requirement. Combining these, we find:

$$\Gamma_m(g)(t) \leq \frac{\epsilon}{4} + \epsilon_1(1 + \theta) + \frac{2\|g\|_\infty}{\delta_1^2}\left(\theta + 2t \cdot \theta + t^2 \cdot \theta\right).$$

Since $t \in D$ is bounded and terms with θ vanish as $m \to \infty$, we conclude:

$$\|\Psi_m(g) - g\|_\infty < \epsilon \quad \text{for all } m \geq \max\{m_1, m_2\}.$$

This completes the proof.

□

3.3 An application of the Bohman–Korovkin theorem

To illustrate the practical utility of the Bohman–Korovkin Theorem, we consider a simple approximation scheme based on piecewise linear interpolation.

Assume $f \in C[0,1]$, and for each $n \in \mathbb{N}$, define an operator $L_n(f)$ as the polygonal (piecewise linear) interpolation of f, constructed using the nodes at $\frac{k}{n}$, where $k = 0, 1, \ldots, n$.

It is straightforward to verify that each operator L_n is linear and positive. Moreover, we observe that:

$$L_n(f_0) = f_0 \quad \text{and} \quad L_n(f_1) = f_1,$$

where $f_0(x) = 1$ and $f_1(x) = x$, since linear functions are preserved under linear interpolation at equally spaced nodes.

Next, we examine the behavior of $L_n(f_2)$, where $f_2(x) = x^2$. On any subinterval $\left[\frac{k-1}{n}, \frac{k}{n}\right]$, the highest difference observed between the quadratic function f_2 and its piecewise linear interpolant is bounded above by:

$$\left| f_2\left(\frac{k}{n}\right) - L_n(f_2)\left(\frac{k}{n}\right) \right| \leq \left(\frac{k}{n}\right)^2 - \left(\frac{k-1}{n}\right)^2 = \frac{2k-1}{n^2}.$$

Taking the maximum over all $k = 1, 2, \ldots, n$, we obtain:

$$\|f_2 - L_n(f_2)\|_\infty \leq \frac{2n-1}{n^2} \leq \frac{2}{n},$$

which tends to zero as $n \to \infty$. Therefore, $L_n(f_2) \to f_2$ uniformly on $[0,1]$.

Having verified convergence on the Korovkin test functions, the Bohman–Korovkin Theorem guarantees that $L_n(f) \to f$ uniformly on $[0,1]$ for any $f \in C[0,1]$.

Chapter 4

Characterization of Best Approximation and Chebyshev Approximation

4.1 Introduction

In many areas of analysis and applied mathematics, it is frequently advantageous to not only approximate a function—but also to approximate it in the most optimal way possible within a given space of functions. This idea leads naturally to the concept of *best approximation*.

Unlike general approximation results such as those provided by the Weierstrass or Korovkin, which ensure that the approximating functions can approach the target function as closely as desired, best approximation focuses on identifying the closest function to a given one, within a specified space.

The notion of best approximation is central in approximation theory, numerical analysis, and functional analysis. In this chapter, we formally define best approximation in normed linear spaces and establish conditions for its existence and uniqueness.

4.2 Best approximation in the uniform norm

Let $C[s,t]$ denote the space of real-valued continuous functions on the interval $[s,t]$, equipped with the uniform (Chebyshev) norm:

$$\|g\|_\infty = \sup_{u\in[s,t]} |g(u)|.$$

For a given $g \in C[s,t]$ and a fixed positive integer m, the task is to identify a polynomial $q_m \in P_m$ satisfying: $\|g - q_m\|_\infty = \inf_{q \in P_m} \|g - q\|_\infty$.

4.3 Chebyshev's problem and error symmetry lemma

In the mid-nineteenth century, Pafnuty Chebyshev initiated the study of uniform polynomial approximation. The focus was on the minimax challenge: given a continuous function g on the interval $[s,t]$, find a polynomial $q \in \mathbb{P}_m$ such that its largest deviation from g is minimized. Specifically, the aim is to determine $q \in \mathbb{P}_m$ such that

$$\max_{u \in [s,t]} |g(u) - q(u)| = \min_{r \in \mathcal{P}_m} \max_{u \in [s,t]} |g(u) - r(u)|.$$

Chebyshev's insight was to show that a best approximating polynomial must equioscillate. We now state a fundamental property of such best approximants.

Lemma 4.3.1. *For a function $g \in C[s,t]$, let $q \in \mathbb{P}_m$ be a closest polynomial that approximates g on $[s,t]$ with respect to the supremum norm. Define the approximation error as*

$$M = \|g - q\|_\infty.$$

Then, the error function $r(u) = g(u) - q(u)$ takes on both values $+M$ and $-M$. In other words, there exist points $u_1, u_2 \in [s,t]$ such that

$$g(u_1) - q(u_1) = M \quad \textit{and} \quad g(u_2) - q(u_2) = -M.$$

Proof. Assume, for contradiction, that the largest error $M = \|g - q\|_\infty$ is achieved with only one sign. Without loss of generality, suppose there exists a point $u_1 \in [s,t]$ where $g(u_1) - q(u_1) = M$, but for all $u \in [s,t]$, we have $g(u) - q(u) > -M$. Define

$$\delta = \min_{u \in [s,t]} \big[(g(u) - q(u)) + M \big].$$

Now construct $w(u) = q(u) + \frac{M+\delta}{2}$, which remains a polynomial of degree at most m. For all $u \in [s,t]$, compute

$$g(u) - w(u) = g(u) - q(u) - \frac{M + \delta}{2}.$$

Since $g(u) - q(u) \leq M$ and $g(u) - q(u) \geq -M + \delta$, it follows that

$$|g(u) - w(u)| \leq \frac{M - \delta}{2} < M.$$

Thus, $\|g - w\|_\infty < M$, which contradicts the optimality of M. This contradiction establishes that the error $g(u) - q(u)$ must achieve both $+M$ and $-M$. □

As a simple consequence, we can determine the best uniform approximation by a constant function.

Corollary 4.3.1. *For a function $g \in C[s,t]$, suppose we approximate g using a constant k. The optimal constant (best uniform approximant in $\mathbb{P}_0$) is given by*

$$k^* = \frac{B + b}{2},$$

where $B = \max_{u \in [s,t]} g(u)$ and $b = \min_{u \in [s,t]} g(u)$. The associated supremum approximation error is

$$D^* = \frac{B - b}{2}.$$

To describe and characterize polynomials of best approximation, we begin by introducing some useful notations and definitions involving continuous functions defined over closed intervals.

Definition 4.3.1. *Let $g \in C[a,b]$,*

1. *A point $x \in [a,b]$ is called a **positive extremal point** if $g(x) = \|g\|_\infty$.*
2. *A point $x \in [a,b]$ is called a **negative extremal point** if $g(x) = -\|g\|_\infty$.*
3. *A sequence of distinct points $a \leq x_0 < x_1 < \cdots < x_n \leq b$ is said to form an **alternating set** for g if $|g(x_i)| = \|g\|_\infty$ for all i and $g(x_i) = -g(x_{i-1})$ for $i = 1, \ldots, n$.*

4.4 Existence of an alternating set

Theorem 4.4.1. *[4] Suppose $g \in C[a,b]$ and $p \in P_n$. If p yields the best uniform approximation of g, then the difference $g - p$ exhibits an alternation at no fewer than $n + 2$ distinct points.*

Proof. Suppose $g \notin P_n$; if it did, then $g = p$ and the statement would be straightforward. Let $E = \|g - p\|_\infty$, which is strictly positive. Define the function $\varphi(x) = g(x) - p(x)$. Since φ is continuous on the interval $[a, b]$, it is also uniformly continuous there. We divide $[a, b]$ into subintervals $a = t_0 < t_1 < \cdots < t_k = b$ such that For every $x, y \in [t_i, t_{i+1}]$, it holds that:

$$|\phi(x) - \phi(y)| < \frac{E}{2}.$$

In any subinterval where $\phi(x) = E$ for some x, the function remains above $E/2$; similarly, where $\phi(x) = -E$, it stays below $-E/2$. Thus, each such subinterval can be classified as either a **(+)** interval or a **(–)** interval, and none can contain both types.

Label these intervals as $I_1, I_2, \ldots, I_m$, alternating in sign, and define $S = \bigcup_{j=1}^{m} I_j$, and $N = [a, b] \setminus S$. Suppose $m < n + 2$. We construct a contradiction.

Between each adjacent pair of oppositely signed intervals, choose a point z_i, and define a polynomial $q(x) = \prod_{i=1}^{m-1}(x - z_i)$. Since $m - 1 \leq n$, we have $q \in \mathcal{P}_n$.

Now, choose $\lambda > 0$ such that $\lambda\|q\|_\infty < \min\left\{E - \varepsilon, \frac{E}{2}\right\}$, where $\varepsilon = \max_{x \in N} |g(x) - p(x)|$.

Define $p^* = p + \lambda q \in \mathcal{P}_n$. For $x \in N$, we get $|g(x) - p^*(x)| \leq \varepsilon + \lambda\|q\|_\infty < E$.

For $x \in S$, since $\phi(x) = g(x) - p(x)$ and $q(x)$ have the same sign, we obtain:

$$|g(x) - p^*(x)| = |g(x) - p(x) - \lambda q(x)| = |\phi(x) - \lambda q(x)| < E.$$

This contradicts the premise that p provides the best approximation. Therefore, it follows that $m \geq n + 2$. □

4.5 Sufficiency of alternating set

Theorem 4.5.1. *[4] Consider $g \in C[a, b]$ and $p \in \mathcal{P}_n$. If the function $g - p$ exhibits alternating deviations in sign at at least $n + 2$ distinct points in $[a, b]$, then p serves as the best uniform approximation to g from the space $\mathcal{P}_n$.*

Proof. Suppose, for the sake of contradiction, that there exists a polynomial $q \in \mathcal{P}_n$ satisfying $\|g - q\|_\infty < \|g - p\|_\infty$. Let $\{x_0, x_1, \ldots, x_{n+1}\}$ denote the set of points where the function $g - p$ alternates in sign.

At each of these points,

$$|g(x_i) - p(x_i)| = \|g - p\|_\infty > \|g - q\|_\infty \geq |g(x_i) - q(x_i)|.$$

It follows that the difference $q(x_i) - p(x_i)$ must share the same sign as $g(x_i) - p(x_i)$, implying that $q - p$ changes sign at least $n + 2$ times. But since $q - p \in \mathcal{P}_n$, it cannot have more than n sign changes unless it is identically zero. Thus, $p = q$, which contradicts our assumption. Hence, p is indeed the best approximation. □

4.6 Uniqueness of best approximation

Theorem 4.6.1. *[4] Let g be continuous on $[a, b]$. Then the best approximation polynomial $p \in \mathcal{P}_n$ to g in the uniform norm is unique.*

Proof. Suppose there exist two distinct best approximations $p, q \in \mathcal{P}_n$ such that:

$$\|g - p\|_\infty = \|g - q\|_\infty = E.$$

Define $r(x) = \frac{p(x)+q(x)}{2} \in \mathcal{P}_n$. Then,

$$g(x) - r(x) = \frac{1}{2}\left[(g(x) - p(x)) + (g(x) - q(x))\right],$$

Hence, $\|g - r\|_\infty \leq E$.

By the existence theorem, $g - r$ must also attain E on an alternating set of at least $n + 2$ points. At each of these points, both $g - p$ and $g - q$ must attain $\pm E$, implying:

$$g(x_i) - p(x_i) = g(x_i) - q(x_i).$$

Thus, $p(x_i) = q(x_i)$ at $n+2$ points. Since $p-q \in \mathcal{P}_n$, it must be the zero polynomial. Therefore, $p = q$, proving uniqueness. □

4.7 Example: Best approximation of sin(4x)

We examine the function $g(x) = \sin(4x)$ on the interval $[-\pi, \pi]$ and seek to determine its optimal polynomial approximation from the space $\mathbb{P}_n$ in the uniform norm.

Initially, consider the zero polynomial $p_0(x) = 0$, which yields a constant approximation to g. The error is then given by $g(x) - p_0(x) = \sin(4x)$. This error function has extrema at equally spaced points across $[-\pi, \pi]$, and it alternates in sign eight times, resulting in eight extremal points where the error achieves its maximum magnitude.

According to the characterization of best approximation, a polynomial $p_n(x)$ is a best approximation to $g(x)$ if and only if the error function $g - p_n$ alternates in sign at least $n + 2$ times.

Hence:

- For $n = 0, 1, \ldots, 6$, the error function alternates at least $n + 2$ times (i.e., 8 times), so $p_n(x) = 0$ is the best approximation in each case.
- For $n = 7$, the condition $n + 2 = 9$ is not met by the 8 sign changes of $\sin(4x)$. Therefore, the constant zero polynomial is no longer optimal, and a better approximating polynomial exists in $\mathcal{P}_7$.

This example illustrates the importance of the alternation criterion in determining whether a given polynomial is the best approximation in the uniform norm.

4.8 Chebyshev's approximation problem

A significant contribution to best approximation theory comes from Pafnuty Chebyshev, who studied the function $g(x) = x^n$ on the interval $[-1, 1]$ and posed this key question:

Which polynomial of degree at most $n - 1$ provides the optimal approximation to x^n on $[-1, 1]$ under the uniform norm?

This problem can be reformulated as: Find $p_{n-1}(x) \in \mathbb{P}_{n-1}$ such that

$$\|x^n - p_{n-1}(x)\|_\infty = \min_{q \in \mathbb{P}_{n-1}} \|x^n - q(x)\|_\infty.$$

In other words, the task is to identify the monic polynomial of degree n that minimizes the uniform norm over $[-1, 1]$.

Chebyshev's construction of the best approximation

We now examine Chebyshev's classical approach to determining the best polynomial approximation to the monomial x^n by a polynomial of degree at most $n-1$ over the interval $[-1,1]$ under the uniform norm.

Step 1: Defining the error function

Let the error function be defined as

$$E(x) = x^n - p(x),$$

where $p(x) \in P_{n-1}$ is the approximating polynomial. Let the uniform norm of the error be

$$M = \|E\|_\infty = \max_{x\in[-1,1]} |E(x)|.$$

Assume that the maximum error is attained at $n+1$ distinct points $x_0, x_1, \ldots, x_n \in [-1,1]$ such that

$$|E(x_i)| = M, \quad \text{and} \quad E(x_{i+1}) = -E(x_i), \quad \text{for } i = 0, 1, \ldots, n-1.$$

This alternating property implies that $E(x)$ has at least n zeros.

Step 2: Critical points and relative extrema

In this step, we analyze the behavior of the error function $E(x) = x^n - p(x)$, where $p(x) \in P_{n-1}$ is the best approximating polynomial. From Step 1, we know that $E(x)$ alternates in sign at $n+1$ points:

$$-1 \le x_0 < x_1 < \cdots < x_n \le 1,$$

and that

$$|E(x_i)| = M = \|E\|_\infty, \quad \text{with} \quad E(x_{i+1}) = -E(x_i).$$

At each interior point $x_i \in (-1,1)$ for $i = 1, \ldots, n-1$, $E(x)$ attains a local extremum. Since $E(x)$ is differentiable, this implies that

$$E'(x_i) = 0 \quad \text{for all } i = 1, \ldots, n-1.$$

Thus, the derivative $E'(x)$, which belongs to P_{n-1}, has exactly $n-1$ real zeros located strictly within the open interval $(-1, 1)$. This structure of $E'(x)$ plays a crucial role in the subsequent development of Chebyshev's characterization.

Step 3: Polynomial structure of the error square

Consider the polynomial

$$M^2 - E(x)^2.$$

This is a polynomial of degree $2n$, and vanishes at each x_i where $|E(x_i)| = M$. Since the error alternates at the points $x_0, x_1, \ldots, x_n$, we have:

- $x_1, x_2, \ldots, x_{n-1}$ are double roots of $M^2 - E(x)^2$,
- x_0 and x_n are simple roots.

Step 4: Relation to the derivative

The square of the derivative $(E'(x))^2 \in P_{2n-2}$. Consider now the polynomial

$$(1 - x^2)(E'(x))^2,$$

which also has degree $2n$, with:

- Double roots at the points $x_1, x_2, \ldots, x_{n-1}$,
- The roots at x_0 and x_n are simple (multiplicity one).

Step 5: Relationship between error and derivative

We now examine the algebraic relationship between the error function $E(x)$ and its derivative $E'(x)$.

From earlier steps, we know that both the functions

$$M^2 - E(x)^2 \quad \text{and} \quad (1 - x^2) \cdot (E'(x))^2$$

are real polynomials having degree $2n$. Moreover, they have the same roots:

- $E(x)$ reaches its maximum and minimum values at $n+1$ alternating points, meaning $M^2 - E(x)^2$ vanishes at each of these points.
- Since $E(x)$ has local extrema at $x_1, x_2, \ldots, x_{n-1}$, its derivative $E'(x)$ vanishes there, and thus $(E'(x))^2$ vanishes with multiplicity two at each of these $n-1$ interior points. The weight function $(1-x^2)$ adds simple zeros at the endpoints $x = -1$ and $x = 1$.

Hence, both $M^2 - E^2(x)$ and $(1-x^2)(E'(x))^2$ vanish at the same set of points with the same multiplicities. Therefore, they must be scalar multiples of each other. That is,

$$M^2 - E(x)^2 = Q \cdot (1-x^2)(E'(x))^2,$$

for some constant $Q > 0$. To establish the value of Q recall that $E(x) = x^n + \cdots$ is a monic polynomial of degree n, and so its derivative $E'(x)$ has leading term nx^{n-1}. This implies that the leading term of $(E'(x))^2$ is n^2x^{2n-2}, and that of $(1-x^2)(E'(x))^2$ is $-n^2x^{2n}$. Thus, the dominant term on the right side is $-Qn^2x^{2n}$, while the leading term of the left-hand side, $M^2 - E(x)^2$, is $-x^{2n}$.

Matching the leading coefficients, we obtain:

$$-Qn^2 = -1 \quad \Rightarrow \quad Q = \frac{1}{n^2}.$$

Substituting back, we get the key identity:

$$M^2 - E(x)^2 = \frac{1}{n^2}(1-x^2)(E'(x))^2.$$

Evaluating the integrals gives:

$$\arccos\left(\frac{E(x)}{M}\right) = n\arccos(x) + Q,$$

so that

$$E(x) = M\cos(n\arccos x + Q).$$

To determine the constant Q, we use the condition $E(-1) = -M$. Then,

$$-M = M\cos(n\arccos(-1) + Q) = M\cos(n\pi + Q).$$

Solving gives:

$$\cos(n\pi + Q) = -1 \Rightarrow n\pi + Q = \pi + 2k\pi \Rightarrow Q = m\pi.$$

Thus, the error function is given by:

$$E(x) = \pm M\cos(n\arccos x).$$

Step 6: Expression using Chebyshev polynomials

Recall that the Chebyshev polynomial of degree n is represented as:

$$T_n(x) = \cos(n\arccos x).$$

Hence, the error function can be written as:

$$E(x) = \pm MT_n(x).$$

Since $T_n(x)$ has leading coefficient 2^{n-1}, and the leading term of $x^n - p(x)$ must match that of x^n, we require:

$$E(x) = 2^{-n+1}T_n(x).$$

Chebyshev's solution to the best approximation of x^n

Let us now express a remarkable result due to Chebyshev, which identifies the polynomial of best uniform approximation to x^n on the interval $[-1, 1]$ by a polynomial of lower degree.

Theorem 4.8.1. *[4] Let $n \geq 1$. Define the polynomial*

$$p(x) = x^n - 2^{-n+1}T_n(x),$$

where $T_n(x)$ is the Chebyshev polynomial of degree n. Then, $p(x) \in \mathcal{P}_{n-1}$, and for every polyno-

mial $q(x) \in \mathcal{P}_{n-1}$*, we have*

$$\max_{|x|\leq 1} |x^n - p(x)| = 2^{-n+1} \leq \max_{|x|\leq 1} |x^n - q(x)|.$$

Proof. First, note that $T_n(x)$ is a polynomial of degree n and it has leading coefficient 2^{n-1}. Therefore, the term $2^{-n+1}T_n(x)$ has degree n and leading coefficient 1, so subtracting it from x^n eliminates the degree n term. Hence, the resulting polynomial

$$p(x) = x^n - 2^{-n+1}T_n(x)$$

has degree at most $n-1$, i.e., $p \in \mathcal{P}_{n-1}$.

Now, define the error function

$$E(x) = x^n - p(x) = 2^{-n+1}T_n(x).$$

Since $|T_n(x)| \leq 1$ for all $x \in [-1, 1]$, we have

$$|E(x)| = |x^n - p(x)| = 2^{-n+1}|T_n(x)| \leq 2^{-n+1}.$$

To verify that this error is in fact the minimum possible, consider the $n+1$ points:

$$x_k = \cos\left(\frac{(n-k)\pi}{n}\right), \quad \text{for } k = 0, 1, \ldots, n.$$

These points lie in $[-1, 1]$ and satisfy $-1 = x_0 < x_1 < \cdots < x_n = 1$.

For each x_k, using the known identity for Chebyshev polynomials:

$$T_n(x_k) = \cos(n \cdot \arccos(x_k)) = \cos(n \cdot \tfrac{(n-k)\pi}{n}) = \cos((n-k)\pi) = (-1)^{n-k},$$

we obtain:

$$x_k^n - p(x_k) = 2^{-n+1}T_n(x_k) = 2^{-n+1}(-1)^{n-k}.$$

Thus, the error function $E(x) = x^n - p(x)$ alternates in sign at the points x_k and reaches the maximum magnitude 2^{-n+1} at each of them. This is an alternating set of $n+1$ points, which satisfies the condition for best approximation. By Chebyshev's Alternation Theorem, This means that $p(x)$ serves as the one and only optimal approximation to x^n from $\mathcal{P}_{n-1}$ in the uniform norm.

Therefore,

$$\max_{|x|\leq 1} |x^n - p(x)| = 2^{-n+1},$$

and for any other polynomial $q \in \mathcal{P}_{n-1}$, we must have

$$\max_{|x|\leq 1} |x^n - q(x)| \geq 2^{-n+1}.$$

This concludes the proof. □

Corollary 4.8.1. *The monic polynomial of degree n that yields the minimal uniform norm on the interval $[a, b]$ is expressed as*

$$p(x) = \frac{(b-a)^n}{2^{2n-1}} T_n\left(\frac{2x-(a+b)}{b-a}\right),$$

where $T_n(x)$ represents the Chebyshev polynomial of the first kind of degree n.

Proof. To map the interval $[a, b]$ to the standard interval $[-1, 1]$, consider the following transformation

$$t = \frac{2x-(a+b)}{b-a},$$

so that $x = \frac{a+b}{2} + \frac{b-a}{2}t$. Let $p(x)$ be a monic polynomial of degree n defined on $[a, b]$. Define a corresponding polynomial on $[-1, 1]$ via the transformation:

$$\tilde{p}(t) = p\left(\frac{a+b}{2} + \frac{b-a}{2}t\right).$$

Then $\tilde{p}(t)$ is a polynomial of degree n on $[-1, 1]$, and we can write

$$\tilde{p}(t) = \left(\frac{b-a}{2}\right)^n \hat{p}(t),$$

where $\hat{p}(t)$ is monic and of degree n. From the theory of Chebyshev polynomials, we know that the monic polynomial of degree n that minimizes the uniform norm on $[-1, 1]$ is given by

$$\hat{p}(t) = 2^{1-n} T_n(t),$$

since $\|T_n\|_\infty = 1$ on $[-1, 1]$. Substituting back, we obtain:

$$\tilde{p}(t) = \left(\frac{b-a}{2}\right)^n \cdot 2^{1-n} T_n(t) = \frac{(b-a)^n}{2^{2n-1}} T_n(t).$$

Finally, expressing t in terms of x, we have:

$$p(x) = \frac{(b-a)^n}{2^{2n-1}} T_n\left(\frac{2x-(a+b)}{b-a}\right).$$

□

Conclusion

In this thesis, we explored the rich landscape of approximation of functions, beginning with the classical *Weierstrass Approximation Theorem*. We first discussed its existential form and then transitioned to *Bernstein's constructive approach*, which not only affirms the theorem but also provides explicit approximating polynomials.

Building upon this, we studied *trigonometric polynomials* within the space $C_{2\pi}$, leading us to the *Weierstrass Second Theorem*. Then, we examined a broader result known as the *Bohman–Korovkin Theorem*, which revealed that uniform convergence of three test functions can be extended to convergence over all continuous functions defined on a compact interval. This theorem offers a powerful and elegant criterion: if a sequence of positive linear operators defined on a compact interval K converges on just three test functions: 1, x, and x^2 then it converges uniformly on the entire space $C(K)$.

Finally, we delved into the theory of *best approximation in the uniform norm*, focusing on the pioneering work of *Pafnuty Chebyshev*. Through his constructive methods and the properties of *Chebyshev polynomials of the first kind*, we identified polynomials that minimize the maximum error.

Overall, this thesis studied foundational theorems with constructive techniques and unifying principles of approximation theory. It highlights how positive linear operators, classical polynomials, and minimax approximations together form the backbone of approximation theory.

References

[1] S. C. Malik and Savita Arora, *Mathematical Analysis*, New Age International Publishers, 2022.

[2] N. L. Carothers, *A Short Course on Approximation Theory*, Bowling Green State University, 2009

[3] S. Serra Capizzano, *Korovkin theorems and linear positive Gram matrix algebra approximations of Toeplitz matrices*, Linear Algebra and its Applications, vol. 284, pp. 307–334, Elsevier, 1998.

[4] M. H. Mudde, *Chebyshev Approximation*, 2017. https://fse.studenttheses.ub.rug.nl/15406/

[5] J. C. Mason and D. C. Handscomb, *Chebyshev Polynomials*, CRC Press, 2002.